

# Variational dynamics in open spacetimes

Jelle P. Boersma

Department of Applied Mathematics, University of Cape Town  
Rondebosch 7700, Cape Town, South Africa

February 7, 2008

## Abstract

We study the effect of non-vanishing surface terms at spatial infinity on the dynamics of a scalar field in an open Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime. Starting from the path-integral formulation of quantum field theory, we argue that classical physics is described by field configurations which extremize the action functional in the space of field configurations for which the variation of the action is well defined. Since these field configurations are not required to vanish outside a bounded domain, there can be a non-vanishing contribution of a surface term to the variation of the action. We then investigate whether this surface term has an effect on the dynamics of the action-extremizing field configurations. This question appears to be surprisingly nontrivial in the case of the open FLRW geometry since surface terms tend to grow as fast as volume terms in the infinite volume limit. We find that surface terms can be important for the dynamics of the field at a classical and quantum level, when there are supercurvature perturbations.

## 1 Introduction

The idea that surface terms can be important when the Lagrangian method is applied to cosmology has been studied earlier in the context of spatially homogeneous but anisotropic models [1] - [4]. In this case, a surface term appears when the Lagrangian is varied with respect to a spatially homogeneous metric perturbation, and the assumption of spatial homogeneity prevents the vanishing of this term when it is evaluated on an arbitrarily distant compact two-surface. In most other cases where the variational approach is applied to cosmology, surface terms are made to vanish trivially by evaluating only variations with respect to variables which vanish outside a bounded domain. The justification for this approach seems to be that one recovers the ‘correct’ field equations, which are the standard Euler-Lagrange equations. In a cosmological context, this way of reasoning can be questioned, both from a theoretical and an observational point of view. From observations, it is not *a priori* clear which

are the correct equations of motion describing the dynamics of fields at length scales larger than the observable universe, and in different cosmological models. From a theoretical point of view, the relation between extremal action fields and classical physics has a natural foundation in quantum field theory. However, field configurations which vanish outside a bounded domain do not play a central role in quantum field theory, and this assumption may be questioned in a cosmological context when the spacetime itself is not bounded.

In this paper we will study this situation by means of an idealized model, which consist of a Klein-Gordon field in both a spatially flat and a spatially open FLRW universe. Our motivation for studying a scalar field stems from the aim to keep our equations simple, and the possible importance of these fields in the description of the early universe. We will concentrate on the open FLRW geometry, since this spacetime has some specific properties which allow surface terms to become important.

One of these properties is that eigenfunctions of the spatial Laplacian occur in two types. First, there are eigenfunctions with eigenvalues exceeding  $\frac{1}{6}$  times the spatial curvature, and these eigenfunctions are complete in the space of square integrable functions [5]. Second, there are eigenfunctions of the spatial Laplacian with eigenvalues between zero and  $\frac{1}{6}$  times the spatial curvature. This last type of eigenfunctions cannot be square integrated, and they are responsible for long-range correlations in a spatially open universe [6]. In spite of the fact that these perturbations cannot be square integrated, they may naturally occur in an open universe which is created in an exponentially expanding false vacuum [7, 8], or they may be generated during preheating [9].

Another important property of the open Friedmann-Lemaître-Robertson-Walker (FLRW) geometry is that a spatial volume and the surface of its boundary grow at the *same* rate when the infinite volume limit is taken. The combination of large boundary surfaces and the presence of long-range correlations in open spacetimes appears to have an effect on the growth of surface terms at spatial infinity in these spacetimes.

Besides the theoretical reasons which make the open FLRW spacetime an interesting object to study, the open FLRW geometry has gained relevance as a model for the observed universe, with observations favoring a relatively small value of the density parameter [10]. Furthermore, progress has been made in describing the creation of an open FLRW universe from an exponentially expanding false vacuum (see, *e.g.*, [11, 12]), and the theory of perturbations in open FLRW spacetimes has been worked out in greater detail [6] - [16].

This paper is structured as follows. In section 2 we discuss the physical relevance of action-extremizing field configurations, and we show that surface terms can contribute to the variation of the action for square-integrable perturbations. In section 4 we decompose the scalar field perturbations in terms of eigenfunctions of the spatial Laplacian, and we discuss the occurrence of supercurvature modes. The dynamics of the extremal action configurations is considered in section 5, and we recover the usual equation of motion for each perturbation component, with an additional source term, which can be expressed in terms of a surface integral which is evaluated at spatial infinity. We show that this source term can be neglected in the case where there are only subcurvature excitations of the scalar field, but it appears to diverge in

the case where there are supercurvature perturbations. Due to this divergence, extremality of the action can only be defined in the restricted phase-space of field perturbations for which surface terms are finite. Depending on how one restricts the phase-space of field perturbations, a nontrivial source term contributes to the equation of motion for the extremal action configurations. In section 6, we consider the quantum correlation function of the scalar field. In the case where there are supercurvature perturbations, it is shown that the action functional is sensitive to degrees of freedom of the scalar field which have zero  $L_2$ -norm. It therefore appears that the correlation function is not well defined, unless one adopts nontrivial constraints on the phase-space of the scalar field, or one needs to include the zero-norm degrees of freedom in the integration over paths.

## 2 The extremal action principle

In this section we briefly review the variational approach to classical field theory. We then use arguments from quantum field theory to motivate a modified form of the variational method in a cosmological context. Surprisingly, it appears that non-local interactions at a classical level can emerge from the underlying quantum theory with a standard expression for the Lagrangian. While our explicit calculations involve only the simple case of a scalar field, our arguments are relevant in a more general field theoretical context, including general relativity. We will come back to this point at the very end of this paper.

One way of describing the dynamics of a classical field is by formulating a field equation. A specific solution of the field equation is determined by the boundary or periodicity conditions which apply to the system. It is of interest to note that the dynamics of the fields which can be observed in nature are described by field equations which act locally, while mathematical consistency does not require this. Hence, the dynamics of classical fields has a local aspect, in the sense that the field equations involve only the field variables and derivatives thereof at each point. Further, a particular classical field configuration is subject to global constraints, which act in the form of boundary or periodicity conditions. The work on this paper started as an attempt to establish whether the local aspects of the dynamics of fields, which is apparent from the structure of the field equations, are fundamental in nature.

In order to gain a deeper insight in the global aspects of the dynamics of fields, a Lagrangian approach appears to be most suitable. In this approach, an action functional is constructed from the field variables over the entire space-time. The dynamics of the field then follows by requiring that the action is extremal in the space of field configurations. Establishing extremality of the action amounts to showing that the action does not vary at first order, for arbitrary infinitesimal perturbations of the field variables. It is essential to note that the field perturbations which are used to ‘test’ the extremality of the action in the classical description, are purely a mathematical construct. Further, we stress that the choice of the action functional is motivated with the aim to recover the field equations, and hence the Lagrangian description has the same physical content as the field equation.

At this point, let us formulate more precisely the question whether the local form of the interactions in nature is fundamental. On the one hand, it is well known that there exist conserved quantities which are related to global symmetries of the action [17] [18]. Although the existence of conserved quantities suggests an underlying global aspect of the dynamics of classical fields, this global aspect is in fact a consequence of applying Gauss's theorem to a four-divergence, which vanishes locally as at each point in our spacetime as a consequence of the field equation. On the other hand, there is the question whether there can be a non-local coupling between physical fields, which acts at the level of the field equation. In particular, one would like to know whether non-local interactions at a classical level can emerge from an underlying quantum theory for which the Lagrangian has the usual local form. In this paper we will focus on this last question.

Let us now consider in some detail how classical field theory arises as a limit of an underlying quantum field theory. According to the Feynman path-integral approach to quantum field theory, the expectation value of an operator  $O$  which acts on a field  $\psi$ , is given by the formal expression,

$$\langle O \rangle = Z^{-1} \int d[\psi] O[\psi] e^{iS[\psi]/\hbar}, \quad (1)$$

where  $S[\psi]$  is the action functional,  $Z$  is a normalization constant, and  $d[\psi]$  is a measure on the space of field configurations (see, *e.g.*, [19]). The integral is evaluated over all field configurations (paths) which are continuous and which satisfy certain initial or periodicity conditions. One should note that there is considerable difficulty involved in making the path-integral well defined, which is due to the fact that typical paths which contribute to the integral are non-differentiable. In our derivation, where we consider a free field, the different degrees of freedom decouple, and one can ignore those degrees of freedom which vary with infinite frequency.

As  $\hbar$  approaches zero in expression (1), the oscillatory behavior of the integrand suggests that the integral is dominated by those field configurations  $\psi$  which are in some sense near to a field configuration  $\psi_0$  which extremizes the action. Since  $\hbar$  is close to zero when expressed in terms of macroscopic units of time and energy, one therefore expects that classical physics is accurately described by an action extremizing field configuration  $\psi_0$ . The essential difference between this classical limit, and the classical theory which we discussed previously, is the fact that in the former case there are physical field perturbations which probe the phase-space nearby an action extremizing configuration, while in the latter case these field perturbations are purely a mathematical construct. As we will show in the following, this difference can give rise to an essentially different expression which describes the dynamics of the classical field.

As is well known, extremality of the action for  $\psi_0$  implies that this configuration satisfies the classical field equations, *provided* that a surface term vanishes for all paths. In a classical variational treatment, surface terms are set to zero trivially by considering only paths which have compact support. However, this restriction on the type of paths does not occur in the sum over paths (1), and it seems natural to consider field configurations  $\psi_0$  which extremize the action for the most general class of paths for which extremality of the action can be

defined.

Indeed, one should note that in a classical treatment of cosmological perturbations one does not normally assume that perturbations must have compact support. However, if one accepts that classical perturbations do not have compact support, then it seems rather unnatural to require that quantum fluctuations about the classical field configurations have compact support. If this would be the case, then there would be a finite distance beyond which there are still classical perturbations while quantum fluctuations vanish. This appears to contradict the Copernican principle, which is commonly adopted in cosmology.

Considering the relation between classical and quantum physics, it should be mentioned that the path-integral approach does not only explain more than a classical approach (*i.e.*, testable quantum effects), but one also needs to assume more than in classical physics (*e.g.*, the existence of a classical regime [20], as well as various infinite subtractions [19]). One might therefore feel that the validity of the path-integral approach is as questionable as the classical variational approach, when it is applied to cosmological situations where it has not been tested. When seen in this light, the classical assumption that field-perturbations are restricted to have compact support is not proven to be wrong, but rather, it represents one possible choice in a more general class of boundary or asymptotic conditions. Whichever point of view one favors, it seems interesting to investigate the implications of relaxing the assumption that field-perturbations must have compact support. We will discuss these implications in the following.

### 3 Scalar field in FLRW geometry

The line element of the FLRW geometry is given by,

$$ds^2 = -dt^2 + a^2(t) [d\chi^2 + c^{-2} \sinh^2 c\chi (d\theta^2 + \sin^2 \theta d\phi^2)], \quad (2)$$

where  $c = \mathbf{R}^+$  for the spatially open geometry, while the spatially flat and closed geometry are obtained by taking the limit  $c \downarrow 0$  or by choosing  $c \in i \times \mathbf{R}^+$  respectively. We will refer to the geometry with the line element (2) as  $\mathcal{M}$ , while a spatial hypersurface of constant time  $t$  is referred to as  $\Sigma$ .

It follows directly from expression (2) that the surface of a spatial sphere of constant radius  $\chi_0$  grows as fast as the three-volume inside the sphere, when one considers the limit where  $\chi_0 \rightarrow \infty$ . One may therefore expect that surface terms can be equally important as volume terms when we take the infinite volume limit in an open universe. This situation is essentially different from the situation in a spatially flat spacetime, where the surface of a spatial sphere of constant radius  $\chi$  grows by one power of  $\chi$  less fast than the three-volume which is contained inside the sphere.

We will consider a scalar field  $\psi$ , which is described by the Lagrangian density

$$\mathcal{L}[\psi] = -\frac{1}{2}\sqrt{-g} (g^{\mu\nu} \partial_\mu \psi \partial_\nu \psi + m^2 \psi \psi), \quad (3)$$

where  $g_{\mu\nu}$  denotes the FLRW metric (2), and  $g = \det(g_{\mu\nu})$ .

We define the action of the  $\psi$ -field as the integral of the Lagrangian density (3) over the entire spacetime,

$$S[\psi] := \int d^4x \mathcal{L}[\psi]. \quad (4)$$

Note that the integral in this expression does not need to converge. This is not necessarily a problem if one is interested in calculating the variation of the action under a change of the field from  $\psi$  to  $\psi + \delta\psi$ , where  $\delta\psi$  is a suitably small ‘test-perturbation’. The question arises which restriction one has to impose on the test-perturbations  $\delta\psi$  so that the first-order variation  $\delta S$  is well defined. The first-order variation of the action (4) follows by the standard procedure of functional derivation,

$$\delta S = \int d^4x \left( \frac{\delta \mathcal{L}}{\delta \psi} \delta \psi + \frac{\delta \mathcal{L}}{\delta \partial_\mu \psi} \delta \partial_\mu \psi \right). \quad (5)$$

By partially integrating equation (5), where the Lagrangian is given by expression (3), we obtain

$$\delta S = \int d^4x \sqrt{-g} [\psi_{;\mu}^{\phantom{\mu}\mu} - m^2 \psi] \delta \psi - \int d^4x \sqrt{-g} (\psi^{;\mu} \delta \psi)_{;\mu}, \quad (6)$$

where a semicolon denotes the covariant derivative.

Provided that the second term on the right-hand side of equation (6) vanishes for nonzero perturbations  $\delta\psi$ , then the condition  $\delta S = 0$  implies the vanishing of the term in brackets, and hence the field equation holds. This is the case when we consider test-perturbations  $\delta\psi \in D$ , where  $D$  is defined as the class of perturbations which are bounded and which have compact support. However, as we mentioned in the beginning of this section, the restriction to test-perturbations  $\delta\psi \in D$  does not follow from known physical principles, when the spacetime itself is non-compact. Let us therefore try to determine the largest class of test-perturbations for which the variation of the action is well defined. For a scalar field  $\psi$ , and a Lagrangian which is bi-linear in the field variable, it is clear that square integrability of  $\delta\psi$  is a necessary condition for the existence of the variation of the action (6), *i.e.*, we require  $\delta\psi \in L_2(\mathcal{M})$ . It is not *a priori* clear whether  $\delta\psi \in L_2(\mathcal{M})$  is a sufficient condition for the existence of the variation of the action (6), and it may be necessary to restrict the type of test-perturbations further to ensure that  $\delta S$  exists. Assuming that we are able to determine the largest class of test-perturbations  $\delta\psi$  for which  $\delta S$  exists, then it remains a question whether there exist field configurations  $\psi_0$  such that  $\delta S$  vanishes for all perturbations  $\delta\psi$  about  $\psi_0$ .

Let us first address the question whether the restriction  $\delta\psi \in L_2(\mathcal{M})$  is sufficient to ensure the existence of  $\delta S$ . The answer to this question is negative, which we show by an example where the contribution of surface terms to  $\delta S$  diverges, while  $\psi$  is a solution of the field equation and  $\delta\psi \in L_2(\mathcal{M})$ . Since we will focus on surface effects at *spatial* infinity, we require that  $\delta\psi$  can be square integrated over a spatial hypersurface of constant time in the geometry (2), *i.e.*,  $\delta\psi \in L_2(\Sigma)$ , while we do not specify the time dependence of  $\delta\psi$ . It is clear from the expression of the line element (2) that a square integrable

test-perturbation  $\delta\psi$  must approach zero faster than  $1/\chi$  in the spatially flat case, and faster than  $e^{-\chi}$  in the spatially open case. A specific example of a square integrable test-perturbation is given by

$$\delta\psi = (1 + \chi)^{-(1+\alpha)} \partial_\chi \psi \quad \text{and} \quad \delta\psi = e^{-(1+\alpha)\chi} \partial_\chi \psi, \quad (7)$$

in the spatially flat and open case respectively, and  $\alpha \in \mathbf{R}^+$ . By substituting expressions (7) for  $\delta\psi$  into equation (6), and using (2), we find

$$\delta S = -4\pi a^{-2}(t) \int dt d\Omega \lim_{\chi \rightarrow \infty} F(\chi) (\partial_\chi \psi)^2, \quad (8)$$

where  $d\Omega$  denotes the volume element on the unit two-sphere, and  $F(\chi) = \chi^{1-\alpha}$  in the spatially flat case, and  $F(\chi) = e^{(1-\alpha)\chi}$  in the spatially open case. Indeed, expression (8) diverges for some values of  $\alpha \in (0, 1]$ , provided that the term  $\partial_\chi \psi$  does not approach to zero as fast as  $F^{-1/2}(\chi)$  in the limit where  $\chi \rightarrow \infty$ . The variation of the action (8) can therefore be arbitrarily large, for  $\delta\psi \in L_2(\Sigma)$ .

Let us now address the question whether there exist configurations of the  $\psi$ -field which extremize the action for all  $\delta\psi \in L_2(\Sigma)$ , in the cosmologically interesting case where  $\psi$  and  $\partial_\chi \psi$  do not vanish at spatial infinity. We show that the answer to this question is negative. We will therefore use a result which is derived in the following, which states that a field configuration which extremizes the action for all  $\delta\psi \in L_2(\Sigma)$  must be a solution of the field equation. We combine this with the result which was derived earlier in this section, which shows that a solution of the field equation for which  $\partial_\chi \psi$  does not approach to zero at spatial infinity, does not extremize the action for all  $\delta\psi \in L_2(\Sigma)$ . Hence, it follows that action extremizing configurations do not exist for  $\delta\psi \in L_2(\Sigma)$  and  $\partial_\chi \psi$  not approaching to zero at infinity.

In deriving the proof above, we assumed that a field configuration which extremizes the action for all  $\delta\psi \in L_2(\Sigma)$  must be a solution of the field equation. In order to proof this, let us recall that for  $\delta\psi \in D$ , *i.e.*, the class of test-perturbations which are bounded and which have compact support, extremality of the action implies that the field equation holds and vice-versa. Configurations which do not satisfy the field equation can therefore not extremize the action for all  $\delta\psi \in D$ , and since  $D \subset L_2(\Sigma)$  these configurations do not extremize the action for all  $\delta\psi \in L_2(\Sigma)$ . Hence it follows that a field configuration which extremizes the action for all  $\delta\psi \in L_2(\Sigma)$  must be a solution of the field equation, which proves our assumption.

The observation that action extremizing configurations do not in general exist for  $\delta\psi \in L_2(\Sigma)$ , implies that the usual identification between classical physics and action extremizing configurations becomes ambiguous when we allow for perturbations which do not fall off sufficiently fast at infinity. There are several ways by which one could try to resolve the problem which is posed by the non-existence of extremal action configurations for test-perturbations  $\delta\psi \in L_2(\Sigma)$ . We will discuss these possible solutions in the following.

First, let us recall that the restriction  $\delta\psi \in L_2(\Sigma)$  was found to be *necessary* to ensure finiteness of  $\delta S$ , but due to the contribution of a surface term to  $\delta S$  this restriction is not *sufficient*. This observation suggests that the class of test-perturbations  $\delta\psi$  should be restricted further, such that  $\delta S$  is finite for

all  $\delta\psi$ . Although finiteness of  $\delta S$  is easily achieved by requiring that the test-perturbations  $\delta\psi$  fall off sufficiently fast, this does not imply that extremal action configurations exist in the space of test-perturbations for which  $\delta S$  is finite. The reason for this is that the existence of extremal action configurations requires that the surface term contribution to  $\delta S$  vanishes completely, which is clearly a stronger restriction on  $\delta\psi$  than the condition that  $\delta S$  is finite. Although one could restrict  $\delta\psi$  to ensure that the surface term contribution to  $\delta S$  vanishes completely, this would be rather add-hoc since it is not shown that this is the only possible restriction on the class of test-perturbations for which extremal action configurations exist.

Instead of restricting the class of test-perturbations, one could also attempt to remove the contribution of surface terms to  $\delta S$  by modifying the Lagrangian density (3). Let us therefore note that the choice of the Lagrangian density is motivated by the fact that one recovers the Klein-Gordon equation, provided that the variation of the action and the surface term in equation (6) vanish. In the classical variational approach, where surface terms are made to vanish by assuming boundary conditions on  $\delta\psi$ , one therefore has the freedom to add a term to the Lagrangian density which has the form of a four-divergence, since the variation of this term equals a vanishing surface term. In this section we questioned the assumption that the surface term in equation (6) vanishes in perturbed flat and open FLRW spacetimes. However, it is conceivable that one can add a four-divergence term to the Lagrangian (3) such that its variation cancels the surface term in equation (6). Indeed, in the context of Hamiltonian cosmology, as well as in quantum cosmology, it appears to be natural to add a surface term to the Einstein-Hilbert action which has the property that its variation cancels an identical term which arises from the variation of the Einstein-Hilbert action [21] - [23].

Let us now consider whether the same possibility exists in the case where we are dealing with a scalar field. We therefore add a generic surface term to the action (4), which has the form

$$S_B[\psi] = \frac{1}{2} \int d^4x \sqrt{-g} B_{;\mu}^\mu, \quad (9)$$

where  $B^\mu = B^\mu[\psi]$ , and then we consider whether the variation of this surface term may cancel the surface term in equation (6). The variation of  $B^\mu$  follows by the method of functional derivation, *i.e.*, treating  $\psi$  and  $\partial_\nu\psi$  as independent variables:

$$\delta B^\mu = \frac{\delta B^\mu}{\delta\psi} \delta\psi + \frac{\delta B^\mu}{\delta\partial_\nu\psi} \delta\partial_\nu\psi, \quad (10)$$

where we used that  $B^\mu$  cannot depend on higher than first-order derivatives of  $\psi$ . It is clear that any dependence of  $B^\mu$  on higher than first-order derivatives of  $\psi$  contributes terms to the variation of the action which are proportional to the variation of higher than first-order derivatives of  $\delta\psi$ . These terms cannot cancel against the surface term in equation (6), which contains at most first-order derivatives of  $\delta\psi$ , although a cancellation was required.

The requirement that the surface term in equation (6) cancels the surface



term which arises from the variation of  $S_B$  results in the conditions

$$\frac{\delta B^\mu}{\delta \psi} = \psi^{;\mu}, \quad \text{and} \quad \frac{\delta B^\mu}{\delta \partial_\nu \psi} = 0, \quad (11)$$

for all  $\mu, \nu$ , and we used expression (10). The first condition in equation (11) constrains  $B^\mu$  to be of the form  $B^\mu = \psi^{;\mu} \psi + c_1$ , where  $c_1$  is a functional which does not depend on  $\psi$ , while the second condition constrains  $B^\mu$  to be a functional which does not depend on  $\partial_\mu \psi$ . Clearly, both requirements are exclusive, and there exists no functional  $B^\mu$  such that the variation of  $S_B$ , (9), cancels the surface term in equation (6). Note, however, that the precise form of the surface term in equation (6) does change by adding a term of the form (9) to the action. Hence, the contribution of a surface term to the variation of the scalar field action (4) appears to be generic, although its precise form is ambiguous. In the following calculation we will retain the surface term which appears in equation (6), which means that we assume  $S_B$  to vanish.

Having considered the possibility to adopt further restrictions on the type of test-perturbations, as well as modifying the action by adding a surface term contribution, we have not found an argument which shows us that we can neglect the contribution of a surface term to the variation of the action. However, taking the surface term in equation (6) seriously confronts us with the problem that field configurations which extremize the action in the space of test-perturbations for which  $\delta S$  is well defined, do not in general exist. It should be noted, however, that the non-existence of action extremizing field configurations does not need to be a problem if one could show that those test-perturbations for which the action functional is not extremal, have a zero phase-space measure in the space of fields  $\psi$ . Indeed, it is clear that paths of the form (8), which yield large surface terms at spatial infinity, are highly special in the sense that the asymptotic behavior of these paths is correlated with the field  $\psi$  about which we expand. Therefore, one expects that these paths occupy a very small amount of phase-space in the space of field configurations in which extremality of the action is considered, and their relevance for the dynamics of the  $\psi$ -field may be negligible. Note, however, that precisely the same argument applies to the case where  $\delta\psi \in D$ , since in this case  $\delta\psi$  is specified to be exactly equal to zero for arbitrarily large radii  $\chi$ . In order to make these considerations quantitative, it is necessary to introduce a measure on the phase-space of the  $\psi$ -field. We will address this problem in the following sections.

## 4 Perturbations in open FLRW

In order to obtain a quantitative description of the space of field configurations of the scalar field  $\psi$ , it is useful to decompose  $\psi$  and test-perturbations  $\delta\psi$  in terms of eigenfunctions of the spatial Laplacian which are complete in the space  $L_2$  of functions which are square integrable on the hypersurfaces  $\Sigma(t)$ . The reason why it is convenient to use eigenfunctions of the spatial Laplacian, is that this operator is present in the expression for the variation of the action (6). When we ignore the surface term, it is therefore clear that each eigenfunction only couples to itself, and the dynamics of each mode is independent of the dynamics of all other modes.

Let  $Q(x)$  be a solution of the Helmholtz equation, *i.e.*,

$$Q_{;i}^i + (k/a)^2 Q = 0, \quad (12)$$

where  $;i$  denotes the covariant derivative with respect to the coordinate  $x^i \in \{r, \theta, \phi\}$  in the geometry (2),  $a = a(t)$  denotes the scale factor, and  $k \in \mathbf{R}^+$ . In the following, we concentrate on the spatially open geometry (2), while we consider the spatially flat spacetime as a limiting case of the spatially open geometry. A basis of solutions of equation (12), which are complete in the space of  $L_2$  functions on  $\Sigma(t)$ , and which factorize in terms of an angular and a radially dependent part, is given by

$$Z_{qlm} = \Pi_{ql}(\chi) Y_{lm}(\theta, \phi), \quad (13)$$

where  $Y_{lm}$  are the standard spherical harmonics on the unit two-sphere, and the radially dependent functions  $\Pi_{ql}(\chi)$  are solutions of the equation

$$\frac{1}{g_2} \frac{\partial}{\partial \chi} g_2 \frac{\partial}{\partial \chi} \Pi_{ql}(\chi) = \left( k^2 - \frac{l(l+1)}{g_2} \right) \Pi_{ql}(\chi), \quad (14)$$

where  $g_2 = c^{-2} \sinh^2 c\chi$ . Equation (14) has solutions of the form,

$$\Pi_{ql}(\chi) = N_{ql} (\sinh c\chi)^l \left( \frac{-1}{\sinh c\chi} \frac{d}{d\chi} \right)^{l+1} \cos(qc\chi), \quad (15)$$

where  $q$  is defined by  $q^2 = k^2/c^2 - 1$ , and

$$N_{ql} := \sqrt{\frac{2}{\pi}} \left[ \prod_{n=0}^l (n^2 + q^2) \right]^{-1/2} \quad (16)$$

is a normalization factor [24, 25]. Notice that the  $q = 0$  mode solves the Helmholtz equation (12) with a *nonzero* eigenvalue equal to  $-c^2/a^2$ , which equals  $\frac{1}{6}$  times the spatial curvature in the geometry (2).

The radial solutions for the spatially flat geometry are obtained by taking the limit  $c \downarrow 0$  in expression (15), keeping  $k$  fixed,

$$\lim_{c \downarrow 0} \Pi_{ql}(\chi) = \sqrt{\frac{2}{\pi}} k j_l(k\chi), \quad (17)$$

where  $j_l$  denotes the spherical Bessel function [30]. From now on, we assume that the spacetime is open, such that  $c \in \mathbf{R}^+$ , and without loss of generality we may set  $c = 1$  in expression (2) by absorbing a factor  $c$  in the definition of the comoving radial coordinate  $\chi$  and by absorbing a factor  $c^{-1}$  in the definition of the scale factor  $a(t)$ .

It follows from expression (15) that the radial functions  $\Pi_{ql}$  can be written as the product of an oscillating factor  $\cos q\chi$  or  $\sin q\chi$ , and a factor which approaches to zero exponentially as  $\sinh^{-1} \chi$  in the limit where  $\chi \rightarrow \infty$ . Since the modes  $Z_{qlm}$  with  $q \in \mathbf{R}^+$  vary at comoving length scales which are typically smaller than the curvature scale which we have set equal to one in the FLRW geometry (2), these modes are called *subcurvature modes*.

There exist solutions of the Helmholtz equation (12) for which  $k^2 \in (0, 1]$ , which corresponds to imaginary values of  $q \in i \times (0, 1]$ . The explicit expression for these modes is still given by equation (15), where the factor  $\cos(q\chi)$  is replaced by  $\cosh(|q|\chi)$ . The modes  $Z_{qlm}$  with  $q \in i \times (0, 1]$  approach to zero as a constant times  $\exp((|q| - 1)\chi)$  in the limit where  $\chi \rightarrow \infty$ , and since they vary at length scales greater than the curvature scale one calls them *supercurvature modes*.

We define the spatial integration operation by

$$\langle f \rangle := \lim_{\epsilon \downarrow 0} \langle f \rangle(\epsilon) \quad (18)$$

where

$$\langle f \rangle(\epsilon) := \int d\Omega \int_0^{1/\epsilon} d\chi \sinh^2(\chi) f. \quad (19)$$

and  $d\Omega^2$  denotes the volume element on the unit two-sphere. The subcurvature modes  $Z_{qlm}(q \in \mathbf{R}^+)$  are orthonormal with respect to spatial integration,

$$\langle Z_{qlm} Z_{q'l'm'} \rangle = \delta(q - q') \delta_{ll'} \delta_{mm'}, \quad (20)$$

and they are known to be complete in the space  $L_2(\Sigma)$  [5], which consists of equivalence classes of functions  $f$  for which  $\langle |f|^2 \rangle$  exists, where we identify functions  $f$  which differ only on a set of Lebesgue measure zero.

For the supercurvature modes, the indefinite integral over the radius in expression (20) does not exist, so that these modes cannot be normalized in the  $L_2(\Sigma)$  sense. Furthermore, expression (20) diverges when only one of the modes  $Z$  corresponds to a supercurvature mode, and  $l = l'$  and  $m = m'$ . Therefore, the supercurvature modes cannot be decomposed in terms of the subcurvature modes. Mechanisms which may be responsible for the generation of supercurvature perturbations in open spacetimes have been investigated in [14, 8].

The  $\psi$ -field may be expanded in terms of the modes  $Z_{qlm}$ ,

$$\psi(x, t) = \psi^-(x, t) + \psi^+(x, t), \quad (21)$$

where

$$\psi^-(x, t) := \sum_{lm} \int_0^\infty dq \psi_{qlm}(t) Z_{qlm}(x), \quad (22)$$

$$\psi^+(x, t) := \sum_{lm} \int_0^i d\bar{q} \psi_{\bar{q}lm}(t) Z_{\bar{q}lm}(x), \quad (23)$$

where  $x = \{\chi, \theta, \phi\}$ , and the integration over  $\bar{q}$  runs along the imaginary axis in the complex  $\bar{q}$ -plane.

An important class of perturbations, which is believed to occur in the early universe, corresponds to the case where the coefficient of each independent mode is chosen according to a Gaussian probability distribution (see, *e.g.*, [26] - [28]). For this type of perturbation, which is called a ‘Gaussian perturbation’ or ‘random-field’, there are no correlations between the coefficients  $\psi_{qlm}$  for different values of  $q, l$ , and  $m$ . The statistical properties of a random-field are

determined by the variance of the Gaussian probability distribution, which we call  $\sigma$ . In the generic case, where  $\sigma$  depends on  $q, l$ , and  $m$ , one cannot determine the variances  $\sigma(q, l, m)$  from a single realization of a random-field, which is determined by the set of coefficients  $\psi_{qlm}$ . Instead, one would need an infinite *ensemble* of random-fields, in order to deduce the statistical properties, *i.e.*, the variances  $\sigma(q, l, m)$ , according to which these random-fields are generated.

Let us now define the *ensemble average* of a functional as the weighted sum of this functional over all random-fields in an ensemble, where the weight factor is given by the probability for each specific random-field to occur. This allows us to define the two-point correlation function of the  $\psi$ -field as the ensemble average of  $\psi(x)$  times  $\psi(x')$ . A random-field  $\psi(x)$  is said to be statistically homogeneous and isotropic when the two-point correlation function is invariant under the group of isometries on  $\Sigma$ , *i.e.*, the group of rotations and spatial translations. Clearly, the two-point correlation function of a statistically homogeneous and isotropic random field can only be a function of a distance measure which is invariant under the group of isometries on  $\Sigma$ , and we can take this distance measure to be the length  $d(x, x')$  of a geodesic which relates the points  $x$  and  $x'$ . It can be shown that statistical homogeneity and isotropy of a random-field  $\psi(x)$  holds if and only if the variances  $\sigma$  do not depend on the labels  $l$  and  $m$  [29].

Although it seems rather artificial to introduce the concept of an ensemble in the context of cosmology, since we can only observe one universe, a physical interpretation of the ensemble average is provided by the property of *ergodicity*. In the context of random-fields, ergodicity is defined as the equivalence of ensemble averaging and spatial averaging, where the spatial average of the two-point correlation function is defined by summing  $\psi(x)$  times  $\psi(x')$  over random sets of points  $x$  and  $x'$  for which the geodesic distance  $d(x, x')$  has a specific value. In the case where  $\Sigma$  is a Euclidean three-space, ergodicity can be proven to hold under fairly weak assumptions [26], but for a hyperbolic three-space no proof seems to be known, while it is usually assumed.

In the following, we will assume a Gaussian statistically homogeneous and isotropic spectrum of subcurvature perturbations. One should note that this type of perturbation cannot be square integrated. This follows by substituting the expansion of  $\psi$ , (22), into the hypersurface integral (18) and using the orthonormality relation (20). The resulting expression contains an indefinite sum over  $l$  and  $m$  of the squared coefficient  $\psi_{qlm}$ , and this sum diverges when the variance  $\sigma(q)$  is nonzero. It is therefore clear that the property of non-square integrability is not specifically related to the presence of supercurvature modes.

## 5 Extremal action dynamics

Let us now calculate the variation of the action (6), which is evaluated over a bounded spatial volume  $V(\chi_0)$ , which we define as those points in the geometry (2) for which  $\chi < \chi_0$ , and then we consider the limit where  $\chi_0 \rightarrow \infty$ . We obtain

$$\delta S = \int dt \lim_{\chi_0 \rightarrow \infty} \left[ a^3 \int d\Omega^2 \int_0^{\chi_0} d\chi \sinh^2 \chi \delta\psi \left( \frac{1}{\sqrt{-g}} \partial_\mu g^{\mu\nu} \sqrt{-g} \partial_\nu - m^2 \right) \psi \right]$$

$$- a \sinh^2 \chi \int d\Omega^2 \delta\psi \partial_\chi \psi \Big|_{\chi=\chi_0}, \quad (24)$$

where  $a = a(t)$ . Using the definition of the integration operation (20), expression (24) can be written in the form,

$$\begin{aligned} \delta S = \int dt \left[ a^3 \left\langle \delta\psi \left( \frac{1}{\sqrt{-g}} \partial_\mu g^{\mu\nu} \sqrt{-g} \partial_\nu - m^2 \right) \psi \right\rangle \right. \\ \left. - a \lim_{\chi_0 \rightarrow \infty} \sinh^2 \chi_0 \int d\Omega \delta\psi \partial_\chi \psi \Big|_{\chi=\chi_0} \right]. \end{aligned} \quad (25)$$

We will consider separately the cases where the expansion of the field  $\psi$  includes only subcurvature modes, and the case where the expansion includes supercurvature modes as well.

### 5.1 Open spacetime with subcurvature perturbations

Let us first consider the case where the field  $\psi$  can be expanded in terms of only subcurvature modes, *i.e.*, we assume that  $\psi_{qlm} = 0$  for all  $q \in i \times (0, 1]$ , so that only the first term in the expansion of the field (21) is nonzero. Equation (25) can then be evaluated separately for each mode, by substituting the expansion (21) into expression (25), and using the orthonormality relation (20). We obtain

$$\begin{aligned} \delta S = \int dt \int dq \sum_{l,m} \delta\psi_{qlm}(t) \left[ a^3 \left( \frac{1}{\sqrt{-g}} \partial_0 g^{00} \sqrt{-g} \partial_0 - a^{-2}(t) k^2 - m^2 \right) \psi_{qlm}(t) \right. \\ \left. - \lim_{\chi_0 \rightarrow \infty} a \sinh^2 \chi_0 \int dq' \psi_{q'lm} \Pi_{ql} \partial_\chi \Pi_{q'l} \Big|_{\chi=\chi_0} \right]. \end{aligned} \quad (26)$$

The requirement that the variation of the action vanishes for nonzero perturbations  $\delta\psi_{qlm}(t)$  implies an equation of motion for each perturbation component  $\psi_{qlm}(t)$ , namely,

$$\left( \frac{1}{\sqrt{-g}} \partial_0 g^{00} \sqrt{-g} \partial_0 - a^{-2} k^2 - m^2 \right) \psi_{qlm}(t) = J_{qlm}, \quad (27)$$

where

$$J_{qlm} := \lim_{\chi_0 \rightarrow \infty} \left[ a^{-2} \sinh^2 \chi_0 \int dq' \psi_{q'lm} \Pi_{ql} \partial_\chi \Pi_{q'l} \Big|_{\chi=\chi_0} \right]. \quad (28)$$

Note that  $J_{qlm}$  acts as a *source term* in equation (27), and this term couples perturbations which have the same angular wave numbers  $l$  and  $m$ . One would like to know whether the limit in expression (28) exists, and whether or not this term can be neglected. In order to answer this question, we need to evaluate the integral over  $q'$  of the distribution  $\psi_{q'lm}$ , which is multiplied by a factor which is of order unity. According to equation (20) and (21), the distribution  $\psi_{q'lm}$  can be defined by,

$$\psi_{q'lm} = \lim_{\epsilon \downarrow 0} \psi_{q'lm}(\epsilon), \quad (29)$$

where

$$\psi_{q'lm}(\epsilon) := \langle Z_{q'lm} \psi \rangle(\epsilon) \quad (30)$$

and the limit  $\epsilon \downarrow 0$  should be evaluated after the integration over  $q'$  is performed. When we integrate over a bounded volume, then the modes  $Z_{qlm}$  are dependent in the sense that their overlap  $\langle Z_{qlm} Z_{q'lm} \rangle(\epsilon)$  is nonzero and of the order of  $\epsilon^{-1}$  for  $q - q'$  of the order of  $\epsilon$ . The number of independent modes in a fixed  $q'$ -interval therefore tends to diverge as  $\epsilon^{-1}$  in the limit where  $\epsilon \downarrow 0$ . In the previous section, we introduced the concept of a Gaussian perturbation. In order to generate a Gaussian perturbation which has an amplitude of order one, the coefficients  $\psi_{qlm}$  in the expansion of the field (22) need to be uncorrelated for values of  $q$  differing more than  $\epsilon$ , while the amplitude of the coefficients must diverge as  $\epsilon^{-\frac{1}{2}}$  when  $\epsilon \downarrow 0$ . The asymptotic behavior of the integral over  $q'$  in expression (28) can therefore be estimated as the sum of  $\epsilon^{-1}$  uncorrelated numbers which are of the order of  $\epsilon^{-1/2}$ , multiplied by a  $q'$ -interval which is of the order of  $\epsilon$ . In the limit where  $\epsilon \downarrow 0$ , the term between brackets in expression (28) will therefore remain of order one, and the expression does not converge. Note, however, that the left-hand side of the equation of motion (27) is proportional to the coefficient  $\psi_{qlm}$ , which diverges as  $\epsilon^{-1/2}$  in the limit where  $\epsilon \downarrow 0$ . We therefore find that the source term on the right-hand side of equation (27) can be neglected in the infinite volume limit, when the perturbations of the field are Gaussian and of the subcurvature type.

## 5.2 Open spacetime with supercurvature perturbations

Let us now attempt to derive an equation of motion for the  $\psi$ -field, in the case where the expansion of the  $\psi$ -field (21) includes supercurvature perturbations.

We may therefore substitute the expansion of the  $\psi$ -field (21) in the expression for the variation of the action (24), which yields,

$$\begin{aligned} \delta S = & \int dt \lim_{\chi_0 \rightarrow \infty} \\ & \times \left[ a^3 \int d\Omega^2 \int_0^{\chi_0} d\chi \sinh^2 \chi \delta\psi \left( \frac{1}{\sqrt{-g}} \partial_\mu g^{\mu\nu} \sqrt{-g} \partial_\nu - m^2 \right) (\psi^- + \psi^+) \right. \\ & \left. - a \int d\Omega^2 \sinh^2 \chi \delta\psi \partial_\chi (\psi^- + \psi^+) \right]_{\chi=\chi_0}. \end{aligned} \quad (31)$$

Using the definition of the integration operation (18), and expression (21), we recover expression (27), with an additional source term which accounts for the coupling between subcurvature and supercurvature perturbations, *i.e.*,

$$\left( \frac{1}{\sqrt{-g}} \partial_0 g^{00} \sqrt{-g} \partial_0 - a^{-2} k^2 - m^2 \right) \psi_{qlm}^-(t) = J_{qlm} + J_{qlm}^+, \quad (32)$$

where  $q \in \mathbf{R}^+$ ,  $J_{qlm}$  is given by expression (28), and

$$J_{qlm}^+ := \lim_{\chi_0 \rightarrow \infty} \int_0^i d\bar{q} \left[ \left( \frac{1}{\sqrt{-g}} \partial_0 g^{00} \sqrt{-g} \partial_0 - a^{-2} k^2 - m^2 \right) \psi_{\bar{q}lm}^+(t) \right]$$

$$\times \int_0^{\chi_0} d\chi \sinh^2 \chi \Pi_{ql} \Pi_{\bar{q}l} + a^{-2} \sinh^2 \chi_0 \psi_{\bar{q}lm}^+ \Pi_{ql} \partial_\chi \Pi_{\bar{q}l} \Big|_{\chi=\chi_0} \Big]. \quad (33)$$

Note that both terms which contribute to expression (33) diverge exponentially in the limit where  $\chi_0 \rightarrow \infty$ , and the limit in this expression does not exist, unless the divergent terms cancel. Let us therefore observe that the two terms at the right-hand side of equation (33) diverge exponentially as  $\exp|\bar{q}\chi|$ , (see section 4), and both terms oscillate due to the radial function  $\Pi_{qlm}$ . A cancellation of the divergent terms in equation (33) requires that both terms oscillate with the same phase. By re-writing equation (33), using,

$$\int_0^{\chi_0} d\chi \sinh \chi \Pi_{ql} \Pi_{\bar{q}l} = \frac{\sinh^2 \chi_0}{q^2 - \bar{q}^2} \Big| \Pi_{ql} \partial_\chi \Pi_{\bar{q}l} - \Pi_{\bar{q}l} \partial_\chi \Pi_{ql} \Big|_{\chi=\chi_0}, \quad (34)$$

one finds that  $J_{qlm}^+$  diverges as the product of an exponential factor  $\exp(|\bar{q}|+1)\chi$ , multiplied by the sum of two terms which oscillate out of phase as  $\Pi_{ql}$  and  $\partial_\chi \Pi_{ql}$ , respectively. Therefore, the right-hand side of equation (32) diverges, and we cannot use this equation to describe the time-evolution of the perturbation component  $\psi_{qlm}(t)$ . Recall that in the absence of supercurvature perturbations, surface terms appeared to give rise to a negligible correction to the equation of motion for each perturbation component  $\psi_{qlm}(t)$ , which followed by requiring that  $\delta S = 0$  for all  $\delta\psi \in L_2(\Sigma)$ . When supercurvature perturbations are present, equations (31) and (32) show that it is precisely a surface term which contributes a divergent term to the variation of the action for all  $\delta\psi \propto Z_{qlm}$ . In this case, the extremal action condition  $\delta S = 0$  cannot be satisfied for all  $\delta\psi \in L_2(\Sigma)$ , irrespectively of the equation of motion which the field satisfies. It is however clear that the condition  $\delta S = 0$  must have solutions when test-perturbations are confined to some subspace of  $L_2(\Sigma)$  for which  $\delta S$  is well defined. We will determine these subspaces in the following.

According to expressions (25) and (21), the surface term which contributes to  $\delta S$  behaves asymptotically as  $\delta\psi$  times a factor  $\sinh^2 \chi \partial_\chi \psi^+$  in the limit where  $\chi \rightarrow \infty$ . The contribution of surface terms to the variation of the action (24) will therefore be finite and convergent, provided that  $\sinh^2 \chi \delta\psi \partial_\chi \psi^+$  converges when  $\chi \rightarrow \infty$ . Let us now *define* the class of test-perturbations  $\{\delta\psi\}_c$  by the requirement that  $\sinh^2 \chi \delta\psi \partial_\chi \psi^+$  converges to a constant  $c \in \mathbf{R}$  when  $\chi \rightarrow \infty$ .

Note that it follows from the definition of  $\{\delta\psi\}_c$  that  $\{\delta\psi\}_c$  contains  $D$ , i.e., the class of functions which are bounded and which have compact support. As is well known, the class of functions  $D$  is infinite dimensional in the sense that there exists a denumerable infinite set of linearly independent basis-functions which is complete in  $D$  [31], and therefore  $\{\delta\psi\}_c$  must be infinite dimensional, for arbitrary  $c \in \mathbf{R}$ . It is therefore not clear whether one class of test-perturbations  $\{\delta\psi\}_c$  for some specific value of  $c \in \mathbf{R}$  dominates in terms of the phase-space which is occupied by these test-perturbations. We will make this statement more precise in the following section, where it is shown that that the classes of test-perturbations  $\{\delta\psi\}_c$ , for different values of  $c \in \mathbf{R}$ , are equivalent up to variations with vanishing  $L_2(\mathcal{M})$ -norm.

Summarizing, we found that the contribution of surface terms to the variation of the action diverges for square integrable field perturbations which do not fall off at a specific rate, depending on the spectrum of supercurvature perturbations. In the presence of supercurvature perturbations, extremality of the action can therefore only be defined with respect to a restricted class of field perturbations. Surface terms contribute a non-trivial source term to the standard Klein-Gordon equation, but the magnitude thereof depends on the choice of the restricted class of test-perturbation with respect to which the action is extremized. The dynamics of the ‘classical’ field configurations therefore remains undetermined, unless one finds a physical argument which constrains the phase-space of the  $\psi$ -field uniquely.

## 6 Quantum correlations

In the previous section we showed that surface terms constrain the phase-space of test-perturbations for which the variation of the Klein-Gordon action is finite, in an open FLRW spacetime with supercurvature perturbations. One may also question whether the nontrivial surface terms which we found have an effect on quantum correlations of the  $\psi$ -field. As is clear from expression (1), the quantum correlation function of the  $\psi$ -field can be expressed as a weighted integral over all continuous field configurations, and the weight factor depends on the source term  $J^+$ , which may be infinite.

The two-point correlation function is given by the formal expression (see, *e.g.*, [19])

$$\tau(x, x') := Z^{-1} \int d[\psi] \psi(x) \psi(x') e^{iS[\psi]/\hbar}, \quad (35)$$

where  $x$  denotes the set of coordinates on  $\mathcal{M}$ .

The standard method to calculate the two-point correlation function is to expand the field  $\psi$ , about some background configuration  $\psi_0$ , in terms of a denumerable complete set of solutions of the four-dimensional Helmholtz equation (see, *e.g.*, [32] for the details involved in this calculation). Since  $L_2(\mathcal{M})$  is known to be separable, there exists a denumerable and complete set of solutions, which we call  $\psi_i$ , and we can choose these solutions to be orthonormal in  $L_2(\mathcal{M})$ . A generic expansion of the field  $\psi$ , about a configuration  $\psi_0$ , takes the form

$$\delta\psi := \psi - \psi_0 = \sum_i a_i \psi_i, \quad (36)$$

where  $a_i \in \mathbf{R}$ . Further, the measure on the space of the field  $\psi$  can be expressed in terms of the coefficients  $a_i$ , *i.e.*,

$$d[\psi] = \prod_i \mu da_i, \quad (37)$$

where  $\mu$  is a normalization constant with the dimension of inverse length, and the indefinite product runs over all values of the label  $i$ .

By substituting the expansions of the field (36) and the measure (37) into the expression for the correlation function (35), the path-integral can be evaluated explicitly. Assuming that there are no nontrivial source terms of the kind which



we discussed in the previous section, then the standard expression for the two-point correlation function follows in terms of the complete set of modes  $\psi_i$ . We will not repeat this calculation here, which can be found, *e.g.*, in [32], but instead we will consider what is the effect on the two-point correlation function (35) when there is a nontrivial source term  $J^+[\psi]$  which contributes to the variation of the action.

Let us now define the set of functions  $\tilde{\psi}_i \in \{\delta\psi\}_0$ , which satisfy the property that the linear span of the modes  $\tilde{\psi}_i$  is dense in  $\{\delta\psi\}_0$ , and the modes  $\psi_i$  are chosen so that they are orthonormal with respect to the  $L_2(\mathcal{M})$ -inner product. We would like to show that the modes  $\tilde{\psi}_i$  are complete in  $L_2(\mathcal{M})$ . Note that the class of functions  $D(\mathcal{M})$ , which are bounded and which have compact support on  $\mathcal{M}$ , is contained in  $\{\delta\psi\}_0$ . But  $D(\mathcal{M})$  is known to be dense in  $L_2(\mathcal{M})$  with the  $L_2(\mathcal{M})$ -norm, and therefore the linear span of the modes  $\tilde{\psi}_i$  must be dense in  $L_2(\mathcal{M})$ . At this point, let us note that the set of functions  $L_2(\mathcal{M})$ , with the  $L_2(\mathcal{M})$ -inner product, form a Hilbert space  $H$ . It is a standard result that a set of functions  $\{\psi_i\}$  is complete in  $H$  when the linear span of the functions  $\psi_i$  is dense in  $H$ , and vice-versa (see, *e.g.*, [33]). This observation implies that the modes  $\tilde{\psi}_i$  are complete in  $L_2(\mathcal{M})$ .

We therefore have two complete and orthonormal sets of functions  $\psi_i$  and  $\tilde{\psi}_i$  in  $L_2(\mathcal{M})$ , and an arbitrary field perturbation  $\delta\psi \in L_2(\mathcal{M})$  can be expressed in terms of the modes  $\tilde{\psi}_i$ , *i.e.*,

$$\delta\psi := \psi - \psi_0 = \sum_i \tilde{a}_i \tilde{\psi}_i. \quad (38)$$

It is simple to show that the transformation which expresses one set of basis functions in terms of the other must be orthogonal. Let us now express the measure  $d[\psi]$ , given by expression (37), in terms of the new set of modes  $\tilde{\psi}_i$ . We obtain,

$$d[\psi] = \prod_i \mu \int d\tilde{a}_i, \quad (39)$$

where we used that the Jacobian of the transformation relating the coefficients  $a_i$  and  $\tilde{a}_i$  equals one when the transformation is orthogonal.

One could expect that the path-integral, evaluated with the measures (37) and (39), gives rise to the same result, since all we have done is to express one complete basis of modes in terms of the other. This observation is not correct. Note that when the path-integral (35) is performed with the measure (39), then the source term  $J^+[\psi]$  vanishes trivially, since the argument  $\psi$  is a linear combination of the modes  $\tilde{\psi}_i$ , and therefore  $\psi \in \{\delta\psi\}_0$ . On the contrary, when the path-integral is performed with the measure (37), then  $\psi$  is a linear combination of the modes  $\psi_i$ , and  $J^+[\psi]$  will generally be nonzero, which follows from the observation that  $J^+[\psi_i]$  diverges for all  $\psi_i$ , as we showed in the previous section.

Let us try to make precise in which sense the expansion of the field in terms of two complete sets of modes (36) and (38) differs. Since both expansions converge to the same limit  $\delta\psi$ , it follows that the *difference* between the two expansions can only be a configuration with zero  $L_2(\mathcal{M})$ -norm. When performing the path-integral (35), using the measures (37) and (39) respectively, we

are integrating over paths in  $L_2(\mathcal{M})$  which may differ by a zero-norm configuration. These zero-norm configurations are precisely the degrees of freedom which give rise to the nontrivial source term  $J^+[\psi]$ . In order to show this, let us recall that  $J^+[\psi_i]$  diverges for all  $\psi_i$ . Since  $J^+[\psi]$  is linear in  $\psi$ , and  $J^+[\tilde{\psi}] = 0$  when  $\tilde{\psi}$  is in the linear span of the modes  $\tilde{\psi}_i$ , it follows that

$$J^+[\psi_i] = J^+[\psi_i - P\psi_i], \quad (40)$$

where  $P\psi_i$  denotes the projection of  $\psi_i$  onto the basis of modes  $\tilde{\psi}_i$ , *i.e.*,

$$P\psi_i := \sum_j \langle \tilde{\psi}_j, \psi_i \rangle \tilde{\psi}_j. \quad (41)$$

But the modes  $\tilde{\psi}_i$  were found to be complete in  $L_2(\mathcal{M})$ , so that  $(1 - P)\psi_i$  must have zero  $L_2(\mathcal{M})$ -norm. The argument of  $J^+$  on the right-hand side of equation (40) has therefore zero  $L_2(\mathcal{M})$ -norm, and therefore this must be the degree of freedom which causes the divergence of the source term. Since the action functional depends on zero-norm degrees of freedom through the term  $J^+[\psi]$ , the expression for the correlation function (35) is under-determined. Recall that the same ambiguity was present when we tried to determine the extremal-action configurations in section 5.2. Although we do not know of a way to resolve this ambiguity, let us consider two different approaches which might work.

First, one can fix the zero-norm degrees of freedom on the basis of a physical or philosophical argument. In practice, this could mean that one sets the source term  $J^+$  equal to zero by restricting the phase-space of the  $\psi$ -field to a dense subset of  $L_2(\mathcal{M})$  for which  $J^+$  vanishes. In order to make this approach better than just guessing, one needs to establish whether specific restrictions on the phase-space of the  $\psi$ -field lead to different predictions, which can be falsified.

As a different approach, one could change the measure on the space of the  $\psi$ -field in order to accommodate the zero-norm degrees of freedom. Again, the problem is that there is no clear guideline for doing so, unless one can show that different choices of measure lead to different observable predictions.

It is illustrative to consider a similar ambiguity which occurs in the definition of the path-integral, when one is dealing with fluctuations at infinitesimal rather than infinite length scales. This ambiguity is related to the fact that typical paths which contribute to the path-integral are non-differentiable. Since the class of smooth paths ( $C^\infty$ ) is dense in the class of continuous paths ( $C^0$ ), the difference between a path in  $C^0$  and the nearest path in  $C^\infty$  must have zero  $L_2(\mathcal{M})$ -norm. As we have seen, the measure (37) does not accommodate these degrees of freedom, and the formal expression is ambiguous on the point of the differentiability of the paths over which we integrate. The action functional is however sensitive to the degree of differentiability of the paths, which is made clear by the fact that the action is generally finite for differentiable paths and infinite for non-differentiable paths. One could try to resolve this ambiguity by simply considering paths in  $C^\infty$ , so that the action functional is well defined, but in this case one can show that the field operators in expression (35) commute trivially, and one does not recover quantum physics [19].

Finally, let us note that similar implications hold for other field theories which are described by an action functional which is non-linear in the field

variable. In particular, it is well known that the Einstein field equations can be derived by varying an action functional, which is given by

$$S[g_{\mu\nu}] = \frac{1}{16\pi G} \int_{\mathcal{M}} R \sqrt{-g}, \quad (42)$$

where  $R$  denotes the Ricci scalar, and we have omitted a possible contribution from matter fields and a cosmological constant. Similar to the case where we considered a scalar field, a contribution of a surface term to the variation of the action does occur. At first-order in the metric perturbation, the contribution of this surface term is given by [34],

$$\delta S[g_{\mu\nu}] = -2 \int_{\partial\mathcal{M}} (\delta K + n^a h^{bc} \delta g_{ab;c}) d\Omega, \quad (43)$$

where  $\delta K$  denotes the variation of the trace of the extrinsic curvature at the boundary  $\partial\mathcal{M}$ , while  $h^{bc}$  and  $n^a$  denote the induced three-metric and the normal to the boundary respectively, and  $d\Omega$  denotes the volume element on  $\partial\mathcal{M}$ . The first term on the right-hand side of equation (43) can be canceled by adding a surface integral of two times the extrinsic curvature  $K$  to the action functional (42) (see also the discussion in section 3). The second term on the right-hand side of equation (43) vanishes when it is evaluated according to a classical variational approach where we set  $\delta g_{ab}$  equal to zero at the boundary  $\partial\mathcal{M}$ , but this term could be of interest in cosmological situations when we do not require that perturbations vanish outside a finite volume.

## 7 Conclusion

We revisited the variational principle in a cosmological context. Starting from the path-integral formulation of quantum physics, we argued that there is a correspondence between classical physics and extremal action fields. The phase-space in which extremality of the action is considered, is not constrained in quantum physics, and we showed that there can be a non-trivial contribution arising from surface terms. We made this problem explicit by considering a scalar field in a perturbed open FLRW spacetime. In the case of an open FLRW spacetime with a Gaussian spectrum of subcurvature perturbations, we found no non-trivial correction to the classical equation of motion. In the case where supercurvature perturbations are present, extremality of the action could only be defined after adopting additional restrictions on the phase-space of the scalar field, but the corresponding equations of motion are ambiguous since they depend on how one restricts the phase-space of the field. We showed that the restricted phase-spaces which yield different physical results, differ by perturbations with vanishing  $L_2$ -norm. This ambiguity is present both at a classical level and a quantum level. We briefly discussed a possible strategy to resolve the ambiguity which is due to perturbations with vanishing  $L_2$ -norm.

## References

- [1] M. MacCallum, in *General relativity, an Einstein centenary survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979).
- [2] G. E. Sneddon, J. Phys. **A 9**, 229 (1976).
- [3] M. MacCallum and A. Taub, Commun. Math. Phys. **25**, 173 (1972).
- [4] M. Ryan, J. Math. Phys. **15**, 812 (1974).
- [5] M. Bander and C. Itzykson, Rev. Mod. Phys. **38**, 346 (1966).
- [6] D. H. Lyth and A. Woszczyna, Phys. Rev. **D 52**, 3338 (1995).
- [7] J. D. Cohn and D. I. Kaiser, gr-qc/9803073.
- [8] D. I. Kaiser, astro-ph/9608025.
- [9] B. Ratra, Phys. Rev. **D 50**, 5252. (1994).
- [10] B. Ratra and P. J. E. Peebles, Astrophys. J. Lett. **432**, L5 (1994).
- [11] M. Bucher, A. S. Goldhaber and N. Turok, Phys. Rev. **D 52**, 3314 (1995).
- [12] M. Sasaki *et al.*, Phys. Lett. B **317**, 510 (1993).
- [13] A. Stebbins and R. R. Caldwell, Phys. Rev. **D 52**, 3248 (1995).
- [14] B. Ratra and P. J. E. Peebles, Phys. Rev. **D 52**, 1837 (1995).
- [15] K. Yamamoto *et al.* Phys. Rev. **D 51**, 2968 (1995).
- [16] Misao Sasaki *et al.*, Phys. Rev. **D 51** 2979 (1995).
- [17] E. Noether, *Invariante Variations Probleme*, Nachr. Ges. Gottingen, 235-257 (1918).
- [18] D. Bleecker, *Gauge Theory and Variational Principles*, Addison-Wesley (1981)..
- [19] R. J. Rivers, in *Path integral methods in quantum field theory*, edited by P. V. Landshoff *et al.* (Cambridge University Press, Cambridge, 1987).
- [20] J. B. Hartle, in *Quantum cosmology and baby universes*, edited by S. Coleman *et al.* (World Scientific, Singapore, 1991).
- [21] T. Regge and C. Teitelboim, Ann. Phys., **88**, 286-318.
- [22] S. W. Hawking, in *General relativity, an Einstein centenary survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979).
- [23] J. B. Hartle and S. W. Hawking, Phys. Rev. **D 28**, 2960 (1983).

- [24] R. Fabbri, I. Guidi and V. Natale, *Astrophys. J.* **257**, 17 (1982).
- [25] J. Garcia-Bellido, *Phys. Rev. D* **55**, 4596 (1995).
- [26] R. J. Adler, *The Geometry of Random Fields* (Wiley, Chichester, 1981).
- [27] S. Karin and H. M. Taylor, *A first course on stochastic processes*, (Academic Press, New York, 1975).
- [28] J. M. Bardeen *et al.*, *Astrophys. J.* **304**, 15 (1986).
- [29] A. M. Yaglom, in *Proceedings of the Fourth Berkeley Symposium Volume II*, edited by J. Neyman (University of California Press, Berkeley, 1961).
- [30] L. F. Abbott and R. K. Schaefer, *Astrophys. J.* **308**, 546 (1986).
- [31] F. G. Friedlander, *Introduction to the theory of distributions* (Cambridge University Press, Cambridge, 1982).
- [32] S. W. Hawking, in *General relativity, an Einstein centenary survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979).
- [33] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That* (Benjamin, New York, 1964).
- [34] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).